Artificial Intelligence, Algorithmic Pricing, and Collusion

Emilio Calvano, Giacomo Calzolari, Vincenzo Denicolò, and Sergio Pastorello

Gabe Sekeres and Finn Ye

April 22, 2025

Cornell University

Background

Motivation

- Algorithms becoming increasingly prevalent in practice
  - German gasoline markets (Assad et al. 2024)
  - Smartphone price discrimination (Kehoe et al. 2020)
- Regulatory questions:
  - How do algorithms get to collusive prices?
  - Can they do so in the absence of active principals?
  - Is algorithmic collusion visibly different than tacit collusion?
- Massive lack of theoretical guarantees for this (see Banchio and Mantegazza (2023); Possnig (2024); Lamba and Zhuk (2025))

Q-Learning Algorithms

- We're familiar with Q-learners
- Specifically, they learn *slowly*. This example has a massive state space and is trying to learn the opponent's policy as part of the state
- Our biggest criticisms are related to this. Specifically:
  - What is the loss in the learning phase?
  - How sensitive are these results to the initialization of the Q-matrix?

Model

Environment

Canonical oligopoly pricing game, with n firms / products and an outside good, where in each period t the demand for good i is

$$q_{i,t} = \frac{e^{\frac{a_i - p_{i,t}}{\mu}}}{\sum_{j=1}^{n} e^{\frac{a_j - p_{j,t}}{\mu}} + e^{\frac{a_0}{\mu}}}$$

where a_i is an index of quality, μ is in index of differentiation, and a_0 is an outside good. Firms choose $p_{i,t}$, and we have exogenous marginal costs c_i. The stage problem is:

$$\max_{p_{i,t}} q_{i,t}(p_{i,t}) \cdot p_{i,t} - q_{i,t}(p_{i,t}) \cdot c_{i,t}$$

This is quasiconcave but does not in general have a nice closed form solution.

Simplified Stage Environment

Assume $n = 2$, $c_i = 1$, $a_i = 2$, $a_0 = 0$, and $\mu = \frac{1}{4}$. Then the stage game reduces to

$$\max_{p_i} \frac{(p_i - 1)e^{8-4p_i}}{e^{8-4p_i} + e^{8-4p_j} + 1}$$

This is strictly concave, and we have that it admits Nash prices

$$p_i^N = p_j^N \approx 1.473$$

and monopoly prices are obtained from setting $n = 1$, where we attain

$$p^M = \frac{5}{4} - \frac{1}{4} W_n(2e^3) \approx 1.925$$

Simplified Stage Environment

We basically have an extension of a Prisoner's Dilemma:

	N	M
N	$(0.22, 0.22)$	$(0.37, 0.12)$
M	$(0.12, 0.37)$	$(0.34, 0.34)$

Since all of the involved functions are continuous and concave, this extends fairly nicely.

So we're making our Q-learners play a repeated Prisoner's Dilemma, and the strategies they learn *should* be similar to canonical repeated PD strategies.